

# *Development of speech fluency over a short period of time: effects of pedagogic intervention*

Article

Accepted Version

Joint article

Tavakoli, P., Campbell, C. and McCormack, J. (2016)  
Development of speech fluency over a short period of time:  
effects of pedagogic intervention. TESOL Quarterly, 50 (2). pp.  
447-471. ISSN 1545-7249 doi:  
<https://doi.org/10.1002/tesq.244> Available at  
<https://centaur.reading.ac.uk/40542/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/tesq.244>

Publisher: TESOL

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online

Tavakoli, Parvaneh; Campbell, Colin & McCormack, Joan

University of Reading

Title: Development of speech fluency over a short period of time: Effects of pedagogic intervention

### **Abstract**

This study investigates the effects of a short-term pedagogic intervention on the development of L2 fluency among learners studying English for Academic purposes (EAP) at a university in the UK. It also examines the interaction between the development of fluency, and complexity and accuracy. Through a pre-test, post-test design, data were collected over a period of four weeks from learners performing monologic tasks. While the Control Group (CG) focused on developing general speaking and listening skills, the Experimental Group (EG) received awareness-raising activities and fluency strategy training in addition to general speaking and listening practice i.e following the syllabus. The data, coded in terms of a range of measures of fluency, accuracy and complexity, were subjected to repeated measures MANOVA, t-tests and correlations. The results indicate that after the intervention, while some fluency gains were achieved by the CG, the EG produced statistically more fluent language demonstrating a faster speech and articulation rate, longer runs and higher phonation time ratios. The significant correlations obtained between measures of accuracy and learners' pauses in the CG suggest that pausing opportunities may have been linked to accuracy. The findings of the study have significant implications for L2 pedagogy, highlighting the effective impact of instruction on the development of fluency.

Key words: L2 fluency, awareness raising, strategy instruction

### **Introduction**

Speech fluency in second language (L2) studies commonly refers to ease or automaticity in the learner speech and is manifested in flow, continuity and smoothness of speech (Segalowitz, 2010; Skehan, 2014). Despite this apparently simplistic definition, speech fluency is a complex phenomenon that interacts with other aspects of performance and encompasses a multitude of linguistic, psycholinguistic and sociolinguistics factors

involved in speech formulation and production (Kormos, 2006; Lennon, 2000; Segalowitz, 2000, 2010). As such, fluency has been perceived as a difficult construct to define (Freed, 2000) and a complex phenomenon to measure (de Jong et al., 2011; Witton-Davies, 2014). Fluency is sometimes used as a holistic concept of language proficiency integrating all different aspects of successful performance, e.g. syntactically complex, lexically diverse and communicatively fluent speech, to name a few characteristics of such performance. However, in a narrower sense of the term, fluency is predominantly associated with the ability to communicate one's intended meaning effortlessly, smoothly and with no or little disruption. Segalowitz (2012, p. 240) suggests that "for most, the qualities that make speech fluent include fast speech, and the relative absence of undue hesitations, pausing, repetition, and repairs."

With the growing interest in using language for communication, influenced by globalization as well as the development of communicative language teaching, achieving speech fluency has become an ultimate goal of language learning for many L2 learners (de Jong & Perfetti, 2011), a major pedagogic objective in various L2 teaching programmes (Seifoori & Vahidi, 2012), and a prime construct in the assessment of oral language proficiency (e.g. ACTFL, 1996; CEFR, 2010). This interest combined with the complexities involved in defining and measuring fluency has given rise to research investigating L2 fluency particularly in task-based language research (Kormos, 2014; Foster, 2001; Foster & Tavakoli, 2009). However, our reading of the relevant literature suggests that not much research has aimed at investigating the effects of pedagogic intervention on the development of speech fluency when learners are studying the language in the target language (TL) context, i.e. when the effects of 'study abroad' (SA) are combined with the effects of fluency-based pedagogic intervention. As such, this article provides a developmental perspective on fluency and attempts to answer the question of whether the L2 fluency of learners studying L2 abroad can be enhanced over a short period of time by employing an effective pedagogic intervention.

### **Theoretical background**

In an earlier definition of fluency, Fillmore (1979) contended that "fluency might simply be the ability to talk at length with few pauses; the ability to fill time with talk; the ability to talk in coherent and semantically dense sentences; the ability to have appropriate things to say in a wide range of contexts; and the ability to be creative and

imaginative in the language use (Fillmore, 1979, p. 51). Fillmore's definition, initially proposed for L1 fluency, has been a starting point for defining fluency in L2 studies, underlining it as a complex and multifaceted construct involving a number of factors and processes in order to produce fluent speech. Freed's (2000) survey of research into fluency reports a continuum ranging from studies of its psychological manifestations, reflections on underlying speech-planning and thinking processes, to studies of speech production, hesitation phenomena, and temporal dimensions of speech. Segalowitz (2000) added to this discussion by highlighting the importance of making a distinction between cognitive and performance aspects of fluency. While cognitive fluency, in Segalowitz's (2000) terms, is concerned with "the efficiency of the operation of the cognitive mechanisms underlying performance", performance fluency considers "the observable speech, fluidity and accuracy of the original performance" (Segalowitz, 2000, p. 202). In a more recent publication Segalowitz (2010) has revisited his conceptualisation of L2 fluency and has suggested that L2 fluency comprises three distinct but inter-related concepts; cognitive, utterance and perceived fluency. While cognitive fluency, in this framework, is defined in terms of the speaker's "ability to efficiently mobilize and integrate the underlying cognitive processes responsible for producing utterances" (Segalowitz, 2010, p. 48), utterance fluency refers to the measurable aspects of fluency such as speed, pausing and hesitation. Perceived fluency refers to the inferences listeners make about someone's cognitive fluency based on their perceptions of how fluent the speaker is, i.e. their utterance fluency. While we have found Segalowitz's definition and the three-way distinction insightful, the scope of the current paper lies in a careful investigation of utterance fluency, and the impact pedagogic intervention would have on it.

### **Development of L2 fluency**

Over the past decades, the majority of studies investigating utterance fluency have primarily drawn on cross-sectional data by focusing on learner performance at one point in time (de Jong et al., 2012; de Jong & Perfetti, 2012; Kormos & Denes, 2004; Mehnert, 1998; Tavakoli, 2011; Tavakoli & Skehan, 2005). This body of research has mainly tried to examine whether proceduralisation and automatization of L2 speech can be facilitated through manipulating task design, performance conditions and/or instructional setting. Studies that have examined the impact of task design and performance conditions on fluency of task performance have reported that specific

aspects of task design, e.g. task structure, have a significant impact on improving the fluency of L2 performance (Ellis & Yuan, 2005; Skehan & Foster, 1996; Tavakoli & Foster, 2008). Similarly, studies that have operationalised L2 learning or performance through adding planning time or task repetition in order to enhance proceduralisation of task performance strongly suggest that such interventions are beneficial in improving L2 fluency at least in the short term (Bygate, 2001; Mehnert, 1998; Seifoori & Vahidi, 2012; Tavakoli & Skehan, 2005). Manipulating task repetition as a means of improving learner fluency, de Jong and Perfetti (2011) asked two groups of learners to repeat the same task three times: one group spoke about the same topic three times whereas the other group spoke about three different topics once each. The results of their study suggest that while task repetition has an impact on gains in different measures of fluency for both groups in the short term, the long term gains in fluency were only observable for the group who spoke about the same topic three times. Although the evidence from studies such as those above, i.e. focusing on cross-sectional data often outside classrooms, firmly supports the hypothesis that L2 instruction can potentially enhance L2 fluency, the main limitation of this kind of research is that it does not provide an insight into the role of pedagogy in improving fluency or the ways in which instruction can influence the development of L2 fluency, an area which many teachers and learners are concerned with outside an assessment setting.

Over the past decades, a number of studies have examined the development of fluency with learners who are based in the TL context. Working with a small sample of four advanced German learners of English over six months of their stay in the UK, Lennon (1990) reported improvement in three aspects of the participants' fluency: speech rate, frequency of filled-pauses and number of pauses per T-units. Given the small sample size and the use of statistical analysis in Lennon's study, it has been difficult to consider the findings as generalizable. Freed, Segalowitz and Dewey (2004) investigated the development of fluency among three groups of learners over a course of one semester in different kinds of settings. The learners studied a) at home (AH), b) on an Immersion summer programme (IM), or c) abroad (SA). The findings suggested that the IM group made significant improvements in terms of the total number of words, length of the longest turn, rate of speech, and speech fluidity. Compared to the AH group, the SA group improved only in terms of speech fluidity but made fewer gains than the IM group. The AH group made no significant gains. Segalowitz and Freed (2004) also found evidence that learner fluency in a SA context improved significantly over a semester in terms of temporal aspects such as speech rate and hesitation aspects of performance like

repetition. Mora and Valls-Ferrer (2012) also reported robust fluency gains for the SA group over three months, while their AH formal instruction group failed to show any improvement in their speech fluency. DeKeyser (2007) concludes that the majority of students studying L2 in the target L2 context over a period of time make measurable progress in their speaking, particularly with regard to the fluency of their speech, but he argues that for a greater degree of proceduralisation and automatization longer periods of time and further opportunities for language use will be needed.

While research evidence suggests that the development of L2 fluency happens in time and as a result of having more opportunities to use language for communicative purposes, preferably in the L2 context, our aim in this study is to examine whether pedagogic intervention can help enhance learner fluency over a limited period of time by raising their awareness about different aspects of fluency and teaching them some fluency strategies. The next section of the paper reviews some of the research into strategy training and awareness raising.

### **Strategy training and awareness raising**

Learner training is a popular means of facilitating L2 acquisition and enhancing learner independence as it is argued that it helps learners make the most of their learning opportunities and promotes their autonomy (Cohen, 2003). Ellis and Shintani (2014) report that strategy training and awareness raising are the two most frequently used types of learner training activities in L2 classrooms and that combining the two is often the preferred approach to learner training (Cohen, 2003). Strategy training involving both explicit instruction and practice has been used in a number of research projects (Guilloteaux & Dornyei, 2008; Lee & Muncie, 2012), and is reported to have positive impacts on L2 learning. Drawing on the findings of research in learner strategies, strategy training claims that successful learners have a wide range of effective strategies at their disposal (Oxford, 1990; Macaro, 2001), and that strategies should be taught in L2 classrooms. Awareness raising, e.g. getting students to 'notice' specific aspects of language, has also been reported to contribute to the development of independent learning and enhance L2 development (Fotos, 2012; Pickering, 2001). In teaching speaking skills direct awareness-raising activities (e.g. engaging learners in description and analysis of examples) are primarily used to enhance accuracy, and indirect awareness-raising activities (e.g. exposing learners to examples) are often

employed to enhance fluency of speech (Kasper, 2006; Barraja-Rohan, 2011). In the present study we aim to use a combination of both kinds of awareness-raising activities to enhance learner fluency. The only study examining learner training in promoting fluency of learner speaking we have noted is Seifoori and Vahidi (2012). Investigating the effects of pre-task planning on fluency and accuracy of performance, Seifoori and Vahidi (2012) trained students over a course of 16 grammar-based conversation sessions to use pre-task planning opportunity as a means of improving their fluency and accuracy. During the 30-minute training sessions over the course of a semester, the learners were exposed to awareness-raising activities and form-focused instruction in order to improve the correct use of different tenses. The findings of this study suggested that the trained online planners produced more fluent speech, but with limited evidence of improvement in their accuracy. Given that the fluency instruction was embedded in the context of form-focused and accuracy-oriented language teaching, and that the training was primarily aimed at a more effective use of planning, it is difficult to determine whether these results have emerged from the effects of successful use of planning time (Tavakoli & Skehan, 2005), proceduralization of the language rules (DeKeyser, 2007), or the development of learner awareness of different aspects of fluency. This is a methodological limitation that future research needs to address.

### **Methodology**

A quasi-experimental design with a pre-test and post-test was used. The experiment was conducted at the language centre of a university in the UK, with a group of 45 students enrolled on an intensive English for Academic Purposes (EAP) course (21 hours of class time per week), aimed at improving their language proficiency to the level required for their academic courses. Although the course had different key component modules, the experiment was conducted in the Listening and Speaking module which typically includes a focus on improving speaking and listening skills, individual learner presentation skills, and participation in class discussions. There were 45 participants in the experiment at the beginning of the study, but the data we are reporting here come from 37 learners, 19 in the Experimental Group (EG) and 18 in the Control Group (CG), who participated in the whole experiment. The participants had all been placed on their course according to their IELTS scores (or equivalent) at the beginning of the course, i.e. four weeks before the experiment started. The language centre's policy is to assign students to groups randomly to ensure a spread of nationalities, L1s and gender. The



EG and CG were also randomly selected from the existing groups of students with IELTS 5-5.5 (or equivalent) in their listening and speaking. The two groups were taught by two different teachers. The learners were at B2 level (CEFR) and had obtained IELTS 5.0 – 5.5 in listening and speaking (or equivalent) before the experiment started. They had a range of different L1s including 11 Arabic, 2 Thai, 7 Chinese, 14 Kazakh and 1 Kurdish, and 2 Indonesian. There were 12 female and 25 male students aged between 22 and 35; in this study we did not consider any implications of nationality, gender or age. It is important to note that like SA students the participants in this study were studying in the target language community. However, unlike SA students, the instruction they received was focused on improving their L2 skills.

### **Procedures and Tasks**

All the participants in both groups received general syllabus-based instruction on listening and speaking activities. The EG further received the pedagogic intervention as part of their syllabus. The intervention in the EG had three aims: a) to raise learner awareness of different aspects of fluency, b) to teach strategies that help improve utterance fluency, and c) to provide opportunities for practicing the strategies in class and independently outside class. The awareness-raising activities included all three aspects of utterance fluency, i.e. speed, breakdown and repair. A range of strategies were taught in the EG including paying attention to personal patterns of frequent pausing, helpful use of single-word and lexical-chunk fillers, and avoiding unnecessary repair moves. The following are examples of the activities used in class to achieve the above-mentioned aims:

- a) Activities to raise awareness of different aspects of fluency
  - students listened to a non-native speaker of English retelling a picture story and evaluated the speaker's fluency in terms of speed, pausing and repair measures
  - students examined the transcript of the picture story retelling and identified where fluency had broken down
- b) Strategies that can be used for improving fluency
  - using lexical fillers (e.g. well) and longer lexical chunks (e.g. let me think) and practising them in conversations
  - avoiding repetitions and hesitations in conversations where possible
- c) Opportunities for practising fluency
  - in class: retelling the picture story that they had listened to in exercise (a)

- at home: retelling another picture story and recording their performance; listening to one's own performance to identify fluency problems; recording their performance of the same task again

The pedagogic intervention lasted for four weeks with two 15-20 minutes instructional sessions per week, and further in-class and follow-up fluency-focused tasks and awareness-raising activities. They had one weekly fluency-training activity as homework (see the “at home” example above). This was then used as a point of discussion or further practice in the next session. The core of activities was aimed at raising awareness about utterance fluency with limited reference to issues related to perceived fluency. Outside the 15-20 minute periods in class, students in the EG had the same opportunities as students in the CG to work on both planned speech, e.g. individual presentations, as well as on spontaneous speech such as unplanned group discussions. The administration of the pedagogic intervention was monitored by the research team and involved careful planning of what to be done during the intervention each week, as well as discussion and feedback sessions with the EG's teacher on a weekly basis.

For the purpose of the study, the participants performed a monologic task twice: the week before the experiment started (Time 1) and the day after the intervention ended (Time 2). They were given 1 minute to plan their monologue and 1 minute to perform it. To avoid any practice effect of listening to someone else performing the task, they were asked to perform it all at the same time. To ensure ecological validity (Eckerth, 2009), the tasks were carefully designed in consultation with the EAP course instructors. The tasks were very similar in structure and in what they required the learners to do, but they had different topics in order to avoid any practice effect. Three main criteria were considered when designing the tasks: a) the tasks were in line with the course objectives, i.e. improving the learners' speaking and listening ability, b) the instructors considered the task as interesting, relevant and at the right level, and c) the task type was familiar but the topics had not been covered on the course before (see the tasks in Appendix 1). It was intended that the nature of the pre- and post-test tasks would allow students to talk fairly easily without having to think too much about content, and were similar enough in design to allow comparison. The tasks were piloted before the intervention started, with minor modifications made in terms of task instructions. Although a counterbalanced design would help reduce the potential effects of task content on performance, this design was not employed in this study as it would have

been considered disruptive to the conduct of the class. All pre and post-test tasks were performed during class time and were digitally recorded.

### **Measuring L2 fluency**

Given its complex and multifaceted nature, speech fluency is known as a difficult construct to measure. L2 fluency is primarily related to how efficiently the speaker is able to mobilize the underlying processes of speech planning and production (Segalowitz, 2010), and as such it is directly affected by the cognitive demands of a task. Fluency interacts with other aspects of performance, e.g. the lexical and syntactic complexity of the speech produced. Performing a cognitively demanding task which may require complex lexis and language structures has been shown to have a significant impact on L2 fluency (Robinson, 2003; Skehan, 2009; Tavakoli & Foster, 2008). In addition, research has provided evidence that utterance fluency is related to a range of individual variations and personal styles (Derwin et al., 2009; de Jong et al., 2012), suggesting that L2 fluency is at least to some extent predictable from an individual's L1 fluency. Research findings have further suggested that different languages may have different fluency characteristics and patterns in terms of the temporal features of speech (de Jong et al., 2012). Following from such findings, Segalowitz (2010) calls for an investigation of L1 fluency when studying L2 fluency. However, this variable would be difficult to examine when investigating L2 fluency in a group of learners with a range of different L1s or in studies involving larger number of participants.

Discussions of which measures are the best indicators of L2 fluency have failed to reach consensus. Skehan (2003) proposed that fluency should be measured in terms of its three main characteristics: a) speed fluency, i.e. speed with which speech is performed, b) breakdown fluency, the pauses and silences that break down the flow of speech, and c) repair fluency, hesitations, repetitions and reformulations that are used to repair speech during the production process. Revisiting this framework, Skehan (2014) suggests that in measuring fluency a distinction should be made between the disturbances made to the flow of speech, e.g. pausing and reformulations, and those made to the speed of speech, e.g. speech rate. This new framework, in effect, groups breakdown and repair fluency measures under *flow*, and distinguishes them from *speed* fluency measures. Skehan (2014) also proposes that composite measures such as phonation time and length of run, i.e. measures that blend *speed* and *flow*, should be considered. Recent

research findings suggest that some measures of fluency are related and, if not chosen carefully, one measure may overlap with others (Kormos, 2006; Skehan, 2014; Tavakoli & Skehan, 2005). One problem in using composite measures, e.g. speech rate, is that since they combine pausing and speed aspects of fluency (de Jong et al., 2012), it would be difficult to identify whether an increase or decrease in fluency has been caused by the pausing or the speed aspect of it. However, as the choice of measures is always guided by the purpose of a study, researchers may choose to use a composite measure such as speech rate if they are more interested in a global measure of fluency, or a non-composite measure such as articulation rate when interested in an individual aspect of fluency, e.g. speed. In a detailed discussion of the best measures of fluency that can both reliably capture L2 utterance fluency and restrict the possible overlap between different measures, Witton-Davies (2014) and Valls-Ferrer (2012) suggest that pause length, pause frequency, pause location, mean length of run, speech and articulation rates, phonation time ratio, and a selection of repair measures are the most reliable measures of utterance fluency. In the current study, in order to see which measures develop more quickly over a short period of time and which are more sensitive to pedagogic intervention, we use a range of measures including both composite and non-composite measures. Following from de Jong et al., (2012) a pause is an unfilled silence of longer than 0.25 a second. The measures of fluency examined in this study are:

- Mean length of run: the mean number of syllables between two pauses
- Mean length of pauses
- Mean number of clause-internal versus clause-external silent pauses
- Mean number of filled pauses, e.g. *em* and *er*
- Repair measures: mean number of partial or complete repetitions, hesitations, false starts and reformulations per minute
- Phonation time ratio: time taken to perform the task (excluding the pauses)
- Articulation rate: mean number of syllables per minute divided by amount of phonation time (excluding pauses)
- Speech rate: mean number of syllables per minutes divided by total time (including pauses)

In task-based language learning research, one approach to defining successful L2 performance is to characterize it as complex, both syntactically and lexically, accurate and fluent. Although the focus of the current study is on the development of fluency, it seems essential to examine learners' performance in full and to investigate the inter-

relationship between different aspects of performance. Whereas studying the development of complexity and accuracy in themselves is an interesting and worthwhile area of research (as discussed below), for the purpose of this study we are interested in accuracy and complexity only in so far as they interact with fluency.

Norris and Ortega (2009, p. 555) call for “more organic and sustainable practices in the measurement of complexity, accuracy and fluency (CAF) in second language production.” They argue that given the challenging task of measuring L2 performance, it is necessary to pay more attention to the ways CAF are measured, defined and interpreted. The challenges involved in this measurement, according to Norris and Ortega (2009), are due to two main reasons. First, since complexity, accuracy and fluency are each a complex system with a number of sub-categories, it is necessary to measure them in a way that a full picture of each construct can be provided. Second, given that CAF is a dynamic system in which the three components interact with one another as they grow and change, it is necessary (albeit a demanding task) to capture the developmental nature of the three measures. To this end, syntactic complexity is often measured through length of units, subordination, coordination and variety. Following on from previous research (Foster & Tavakoli, 2009; Norris & Ortega, 2009) and in order to ensure that complexity is measured efficiently, we have chosen one measure of subordination, i.e. mean number of clauses per AS-unit (Foster, Tonkyn & Wigglesworth, 2000), and one measure of length, i.e. mean number of words per AS-unit. Lexical complexity is not measured in the current study as previous research (Foster & Tavakoli, 2009; Tavakoli & Foster, 2008) has shown that different task content and design may elicit very different sets of lexical items. For accuracy, we have used a global measure, i.e. percentage of error-free clauses, and a local measure, i.e. percentage of correct verb use. The correct use of verb examined the accuracy of use in terms of tense, person, aspect and lexical choice.

## **Data Analysis**

The learners' task performances were digitally recorded, transcribed and coded for a range of fluency measures, i.e. mean length of run, mean length of pauses, mean number of clause-internal versus clause-external silent pauses, mean number of filled pauses, repair measures, phonation time ratio, articulation rate, and speech rate. Although they were asked to speak for one minute, many participants spoke longer for each task. For the purpose of the analysis, all the measures are based on the first 60 seconds of their performance, and the use of appropriate software has ensured a high

degree of precision and reliability. In the current study PRAAT (Boersma & Weenink, 2013) software was used to measure temporal aspects of fluency e.g. phonation time, length of pause or articulation rate. Intra-rater reliability was used for a 10% sample of the data and coefficient measures of above 95% were achieved. For the rest of the measures, e.g. the number of filled pauses, complexity and accuracy, coding was done manually with a researcher reading through the text and coding the transcripts. To ensure reliability of the coding of these measures, initially a 10% sample of the data were coded by a second researcher. This was repeated for a second and sometimes a third time until a 90% inter-rater reliability was obtained.

The following research questions guided the current study:

1. Does L2 fluency in monologic task performance improve over a period of four weeks when learners study L2 abroad on an intensive EAP course?
2. Does pedagogic intervention (awareness-raising and strategy training) have an impact on L2 learners' speech fluency, (in monologic task performance), over a limited period of four weeks?
3. Which aspects of fluency are more sensitive to learner training?
4. Does the development of L2 fluency interact with accuracy and/or complexity of learner performance?

## **Results**

### **Development of Fluency over a Four-week Period**

All data were tabulated and analyzed using SPSS version 18. Table 1 below shows the descriptive statistics for all measures of fluency at Time 1 and Time 2 of the study. The figures indicate that some aspects of learner fluency generally improved over the period of four weeks. Overall, at Time 2 the learners were able to produce longer runs, had higher articulation and speech rates and fewer pauses at clause boundaries. An improvement in reducing the number of repair measures was only observed for EG.

INSERT TABLE 1 HERE

### **Effects of Pedagogic Intervention**

First, a number of independent *t*-tests were run to find out whether EG and CG were different in terms of fluency at the beginning of the experiment. The results indicated

that there were no statistically meaningful differences between the two groups. The only measure reaching a significant level was the mean number of filled pauses in which EG had produced more filled pauses ( $t = 2.30$ ,  $p < .03$ ). While this difference may suggest the learners in the CG were using fillers more frequently at the start of the intervention, we considered this was only a minor indication of group differences and would not violate the assumption of homogeneity of groups before the intervention.

We then ran a repeated-measures multivariate analysis of variance (MANOVA) to investigate whether there were statistically meaningful differences between the two groups' fluency measures at the two points of the data collection. The dependent variables were the different measures of fluency (see Measuring L2 Fluency section above). The independent variables were Time (Time 1 vs. Time 2) and Group (EG vs. CG). All the preliminary assumptions of normality and linearity were checked and no violations were noted. The analysis confirmed the overall effect of time and intervention on the dependent variables, indicating two statistically significant differences in the data, one for Time (Wilks' Lambda = .47;  $F = 6.99$ ,  $p = .001$ ;  $\eta^2 = .53$ ) and one for the interaction between Time and Group (Wilks' Lambda = .66;  $F = 3.08$ ,  $p = .008$ ;  $\eta^2 = .32$ ). We then ran post-hoc  $t$ -tests to identify where the significant differences were located. The non-significant results obtained for Kolmogorov-Smirnov tests confirmed the normality of the distribution for all the measures. In order to avoid running an increased risk of Type 1 error, i.e. the risk of having some spurious positive alpha levels, a Bonferroni-adjusted alpha level of .025 was considered for the paired  $t$ -tests. The results of the  $t$ -tests comparing fluency gain scores from Time 1 to Time 2 are presented in Table 2 below. Where significant results are achieved, effect size measures are provided.

INSERT TABLE 2 HERE

As shown in Table 2, there were no statistically significant differences for gains in fluency measures in the CG. The only measure that approached a significant level was the mean number of filled pauses. Incidentally, the results showed that the learners in the CG produced more filled pauses at Time 2, though at both Time 1 and Time 2 the CG produced fewer filled pauses than the EG. For EG,  $t$ -tests were significant for four of the nine fluency measures: these are  $p = .017$  for mean length of run,  $p = .016$  for phonation time ratio,  $p < .001$  for articulation rate, and  $p < .001$  for speech rate. Based on Cohen's (1988) interpretation, the effect size figures are small for mean length of run

( $\eta^2=.18$ ), phonation time ( $\eta^2=.11$ ) and articulation rate ( $\eta^2=.26$ ), but a medium effect size is observed for speech rate ( $\eta^2=.40$ ). Although the result for the mean number of repair measures in EG failed to reach a significant level, its effect size was noticeable ( $\eta^2=.11$ ). The figures for effect size are important as they indicate the amount of the total variance in fluency measures that is predictable from the effects of the intervention (Tabachnic & Fidell, 1996).

### **Relationship between Fluency, Accuracy and Complexity**

We then examined correlations between the different fluency measures, and explored possible associations with accuracy and complexity measures. Similar to the findings of previous research (Mora & Valls-Ferrer, 2012; Witton-Davies, 2014), many of the fluency measures in the current study correlated at a significant level ( $p<.05$ ) with one another. Previous research (Tavakoli & Skehan, 2005) suggests that a number of distinct factors represent fluency, but the discussion of how many factors there are or how distinct they are goes beyond the scope of this paper. In the data analysis, speech rate was the measure which most often correlated with other fluency measures. Table 3 below shows the correlations between speech rate and other measures of fluency at Time 1 and Time 2.

INSERT TABLE 3 HERE

Table 4 below shows the descriptive statistics for the two accuracy and two complexity measures in this study. Our findings show that means increased on most measures over the period of four weeks for both groups, with all learners producing more error-free and syntactically more complex language at Time 2. We found that the two measures of accuracy correlated with each other (Time 1:  $r=.521$ ,  $p<.001^{**}$ ; Time 2:  $r=.356$ ,  $p<.03^{*}$ ), as did the two measures of complexity (Time 1:  $r=.604$ ,  $p<.001^{**}$ ; Time 2:  $r=.684$ ,  $p<.001^{**}$ ). No significant correlations were observed between the accuracy and complexity measures.

INSERT TABLE 4 HERE

As the means between the two groups did not appear very different, we used independent *t*-tests to see if the two groups' performances were comparable in terms of accuracy and complexity at Time 1. The results of the *t*-tests showed that the two



groups' performances were not statistically different at Time 1. Then, a number of paired sample *t*-tests were run to examine if the learners' performances at Time 2 were more accurate and/or more complex than those at Time 1. Although the two groups had positive gains in their accuracy and complexity during the four-week period, only one of the gains reached a statistically significant level, i.e. the percentage of error-free clauses in Control Group ( $t = 2.81, p < .01^*$ ). Given the non-significant results for gains in other accuracy and complexity measures, this finding is interesting. The analyses show that while the EG produced more fluent language at Time 2, the CG showed some significant improvement in their accuracy of use of verbs. The lack of a significant improvement in the measure of accuracy for the EG will be discussed later.

Our interest in examining the learners' accuracy and complexity of performance focused on exploring whether the development of fluency interacted with the development of accuracy and complexity. In order to see if there was a relationship between measures of fluency, accuracy and complexity after the intervention, a number of Pearson correlation analyses were run for each of the two groups. First we ran separate Pearson correlation analyses for these measures in the CG and the EG at Time 1. The results indicated that there were no statistically significant correlations between the measures of fluency, and accuracy and complexity for either group at Time 1. Then the same analyses were run for Time 2 measures in each group. The results indicated that while there were no statistically significant correlations between the different fluency measures and the measures of accuracy and complexity in the EG, in the CG the two measures of accuracy correlated positively with the measures for number of pauses. The percentage of error-free clauses correlated with mean number of clause-external pauses ( $r = .584, p < .01$ ), and the percentage of correct verb use correlated with mean number of clause-internal pauses ( $r = .533, p < .01$ ). The significant correlations imply that in the CG the higher the number of pauses, whether mid-clause or end-clause, the more accurate the learners' performance. One way to interpret this finding is that while fluency training encouraged the EG to focus on different aspects of fluency, e.g. to hesitate less frequently, the CG may have used the pausing opportunity to focus on form and produce more accurate performance. This kind of monitored fluency will be discussed later.

## Discussion

The prime aim of the current study was to provide insight into the development of L2 fluency over a short period of time in an instructional SA setting. In response to Research Question 1, the first important finding to report here is that a number of aspects of speed fluency, i.e. articulation and speech rates as well as length of run improved over relatively short period of four weeks. This finding is in line with previous research (Freed et al., 2004; Mora & Valls-Ferrer, 2012; Segalowitz & Dewey, 2004) that reports gains in fluency over the course of a semester in an SA context. However, this finding is different from previous research for two inter-related reasons. First, the L2 fluency development is observed over a shorter period of time, i.e. four weeks. Second, we expected some changes from all learners since they were enrolled on an EAP course aimed at improving their speaking and listening skills, and as such they all received intensive L2 instruction (21 hours per week) over the 4-week period. Explaining whether this improvement is the result of teaching and learning provided on all components of the course, the learners' need to communicate with other L2 speakers inside and outside the classroom, being immersed in the L2 during their stay, or a combination of these factors is beyond the scope of the current study.

With regard to Research Question 2 and the impact of pedagogic intervention, it is clear from the results that compared to the relatively small fluency gains in the CG, the gains in length of run, articulation and speech rates and phonation time ratio in the EG from Time 1 to Time 2 reached a statistically significant level. The noticeable effect sizes for these measures (Cohen, 1998) suggest that the magnitude of the difference between the two groups' fluency gains may have been caused by the pedagogic intervention. In explaining the effects of the pedagogic intervention on L2 fluency in our study, one way to interpret the findings is to argue that providing effective instruction and creating opportunities for practice facilitated the process of proceduralization of learner interlanguage which might have contributed to the learners' preparedness for developing a degree of automatization in their performance (DeKeyser, 2001; 2007; Segalowitz, 2010). One observation that we have made is that the pedagogic intervention used in this study helped learners improve speech fluency in terms of successfully managing the discourse flow. Whether this can be interpreted as an aspect of proceduralization remains a question for future research to investigate. In line with Anderson's (1983) adaptive control of thought theory (ACT), it is also possible to argue that automatization, i.e. a process of development from conscious, controlled and often slow processing of declarative knowledge to a more rapid, effortless and attention-free processing of language, was facilitated in the EG by providing the learners with

communicatively oriented declarative knowledge, e.g. the awareness-raising activities, and by creating useful opportunities for practice, e.g. fluency strategy training. Whether awareness-raising activities provide learners with declarative knowledge, metalinguistic awareness or both opens up an interesting discussion that is beyond the scope of the present paper. While the design of the current study did not allow for the differential effects of awareness-raising activities and fluency strategy training to be examined, we call for further research to look into these variables separately and to evaluate their impact on different stages of the automatization of L2 fluency.

As for Research Question 3, the fluency gains imply that certain aspects of fluency, e.g. some speed measures, lend themselves well to short-term intervention in the form of awareness-raising and strategy training practiced during the intervention.

Interestingly, in addition to improving their speed fluency, the EG also improved in terms of repair fluency measures. Although the EG's gains in repair measures did not reach a statistically significant level, the fewer repair measures produced by the EG suggests that the pedagogic intervention may have encouraged learners to monitor their use of the repetitions, false starts, reformulations and hesitations during performance. Further research is needed to investigate development and use of repair measures in more depth. On the other hand, the results imply that the development of breakdown fluency, i.e. silence and pausing, is slower and less sensitive to pedagogic intervention. The fact that many of the learners produce more pauses at Time 2, either at clause internal or clause external locations, implies that pausing is a more tenacious construct, perhaps an inherent cognitive or articulatory characteristic of the language production process. In line with previous research (Kormos, 2006), our findings imply that both repair measures and pauses act as monitoring processes during speech production, with the former being an overt-monitoring and the latter a covert-monitoring process. These results also draw our attention to recent research findings that suggest L2 fluency is at least to some extent a function of L1 fluency and personal styles (de Jong, et al., 2012). Given that the findings of de Jong et al.'s study (2012, p. 21) suggest that "all fluency measures could, to a certain extent, be predicted on the basis of L1 fluency behaviour", it seems necessary to examine learners' L1 fluency when studying their L2 fluency. Further research is also needed to examine the development of pauses, their different locations, and how they improve over time in individual learners, as they may indicate what underlying psycholinguistic processes are involved in speech production.

Answering Research Question 4, the results also indicated that the accuracy and syntactic complexity of learner performance showed some limited improvement over a four-week period. Interestingly, the only measure that reached a statistically significant level was the percentage of error-free clauses in the CG. The correlation analyses further indicated that in the CG the accuracy measures correlated with the number of pauses, i.e. the learners who paused more frequently at clause boundaries produced more accurate language ( $r = .584, p < .01$ ), and those who paused more frequently in mid-clause positions, produced more correct verb forms ( $r = .533, p < .01$ ). While the learners in the CG appear to have used pausing opportunities to improve their accuracy at Time 2, there is no statistical evidence to suggest that accuracy and fluency were related in the CG at Time 1, or in the EG at Times 1 and 2. This implies that pausing opportunities are linked with the monitoring processes of speech production (Kormos, 2006; Tavakoli, 2011), i.e. pausing provides an opportunity for monitoring online speech production, possibly in an attempt to produce more accurate language, rather than build up more fluent language. Although this finding may imply the possibility of a trade-off between fluency and accuracy (Skehan, 1998, 2009), further research is needed to investigate whether directing learners' attention towards one aspect of performance would have a negative impact on other aspects of their performance.

In line with previous research on L2 fluency (Kormos, 2006; Mora & Valls-Ferrer, 2012; Tavakoli & Skehan, 2005; Witton-Davies, 2014), the current study confirms that measures of fluency of each aspect, i.e. break-down, speed and repair measures (see earlier section on *Measuring L2 utterance fluency*), correlate with one another at a significant level. While this finding implies that all the measures of fluency used in this research reliably capture “the intersubject variation in oral performance in the temporal domain” (Mora & Valls-Ferrer, 2012, p. 625), we argue that future research may need to focus on identifying the best representatives of each aspect of fluency, allowing for comparative and replication studies to be carried out.

As this study has focused on utterance fluency (Segalowitz, 2010), we do not have evidence to claim whether cognitive fluency has also developed. We assume that the learners found performing L2 speaking tasks and communicating in L2 less demanding after the four weeks due to the instruction they received on their course, the interactions they had with other speakers and the opportunities for using and practicing their linguistic resources inside and outside the classroom. To ascertain in what ways cognitive fluency develops, further research is required. This can be done by using different task types and designs (Tavakoli, forthcoming), operationalising the degree of

cognitive complexity (Robinson, 2007; Skehan, 2009) inherent in a task, controlling for individual differences such as working memory (Wright, 2013), and monitoring the external factors that affect interlanguage development, e.g. amount of exposure.

## **Conclusions**

This study has investigated whether a short-term intervention in an instructed SA setting could have an effect on the development of utterance fluency. The result that all students demonstrated some fluency gains over a four-week period is perhaps not surprising, given that similar results have been shown in SA studies, albeit studies of longer duration. Mora and Valls-Ferrer (2012, p. 611) argue that the conditions of the SA context promote automatization of already proceduralized linguistic knowledge as they provide opportunities for meaningful interaction and rich exposure to L2 input. We would like to add to this argument by highlighting the effectiveness of classroom-based learner training that promotes learner awareness and their independent practice. We agree with DeKeyser (2007) that automatization is one of the most problematic stages of acquisition in the classroom context. However, a key finding of the current study is that although the classroom context often provides limited and insufficient opportunities for L2 practice, tailor-made training aimed at improving fluency can have short-term positive effects. Whether these effects are because of proceduralization of language rules or an indication of some limited automatization is beyond the scope of the current study.

The findings of the study confirm the hypothesis that the effects of studying L2 in the TL context combined with effective pedagogic intervention can provide conditions conducive to the development of speech fluency. The findings also clearly suggest that speed (e.g. articulation rate) and composite measures (e.g. mean length of run) seem to develop faster both over time and/or as a result of intervention. These findings have important implications for different dimensions of L2 pedagogy from teacher training to materials development. Although the current study provides clear longitudinal evidence on the development of fluency through instruction, a delayed post-test would have provided an opportunity to examine whether these effects were sustained in time. We were unable to collect specific information about the learners' lives outside the class due to practical considerations. However, based on the course structure and content, we know that the learners had ample exposure to the L2 at the university, if not beyond.

## Acknowledgement

This study was supported by research funding from the Centre for Literacy and Multilingualism (CeLM) at the University of Reading. We would like to thank Anne-Marie Hunter for her invaluable help with the Praat analysis, and the three anonymous reviewers for their constructive and useful feedback on a previous draft of the article.

## Authors

Parvaneh Tavakoli is associate professor in Applied Linguistics and Second Language Acquisition at the Department of English Language and Applied Linguistics at the University of Reading.

Colin Campbell is associate professor in Second Language Education and school coordinator of teacher development at the International Study and Language Institute at the University of Reading.

Joan McCormack is programme director and lecturer in EAP at the International Study and Language Institute at the University of Reading.

## References

- ACTFL (2014). American Council on the Teaching of Foreign Languages. Retrieved 1<sup>st</sup> November 2014 from <http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>
- Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Barraja-Rohan, A-M. (2011). Using conversation analysis in the second language classroom to teach interactional competence. *Language Teaching Research*, 15(4), 479–507. doi: 10.1177/1362168811412878
- Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer. Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>
- Bygate, M. (2001). Effects of task repetition on the structure and control of language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Task-based learning: language teaching, learning and assessment* (pp. 23–48). London: Longman.
- Cohen, A. (2003). The learner's side of foreign language learning: Where do styles, strategies, and tasks meet? *International Review of Applied Linguistics in Language Teaching*, 41(4), 279–291. doi:10.1515/iral.2003.013
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

- Council of Europe. (2014). The Common European Framework. Retrieved 10 February 2014 from [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf)
- De Jong, N., Groenhout, R., Schoonen, R., & Hulstijn, J. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics* 36(2), 223-243. doi:10.1017/S0142716413000210
- De Jong, N., & Perfetti, C. A. (2011). Fluency Training in the ESL Classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 533–568. doi:10.1111/j.1467-9922.2010.00620.x
- De Jong, N., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J. (2012). facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5–34. doi:10.1017/S0272263111000489
- DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). New York: Cambridge University Press.
- DeKeyser, R. (2007). Situating the concept of practice. In R. DeKeyser (Ed.), *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1–18). New York: Cambridge University Press.
- Eckerth, J. (2008). Negotiated interaction in the L2 classroom. *Language Teaching*, 42(01), 109. doi:10.1017/S0261444808005442
- Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. London: Routledge.
- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–192). Amsterdam: John Benjamins.
- Fillmore, C. J. (1979). On Fluency. In C. J. Fillmore, D. Kempler, & W. S. Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85–102). New York: Academic Press.
- Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks : second language learning, teaching, and testing* (pp. 75–93.). Harlow, England ; New York: Longman.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299–323. doi: 10.1017/S0272263100015047
- Foster, P., Tonkyn, A. & Wigglesworth, G. (2000). Measuring spoken language. *Applied Linguistics* 21(3): 354-375. doi: 10.1093/applin/21.3.354

- Foster, P. & P. Tavakoli (2009). Native speakers and task performance: Comparing effects on complexity, fluency and lexical diversity. *Language Learning*, 59(4): 866-896. doi: 10.1111/j.1467-9922.2009.00528.
- Fotos, S. S. (2012). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28(2), 323-351. doi:10.2307/3587436
- Freed, B. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 243-265). Ann Arbor: University of Michigan Press.
- Freed, B., Segalowitz, N., & Dewey, D. (2004). Context of learning and second language fluency in French: comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(02), 275-301. doi:10.1017/S0272263104262064
- Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly*, 42(1), 55-77. doi:10.2307/40264425
- Kasper, G. (2006). Beyond repair: Conversation analysis as an approach to SLA. *AILA Review*, 19, 83-99. doi:10.1075/aila.19.07kas
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145-164. doi:10.1016/j.system.2004.01.001
- Lee, S. H., & Muncie, J. (2012). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly*, 40(2), 295-320. doi:10.2307/40264524
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417. doi:10.1111/j.1467-1770.1990.tb00669.x
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggensbach (Ed.), *Perspectives on fluency* (pp. 25-42). Ann Arbor: The University of Michigan Press.
- Macaro, E. (2001). *Learning strategies in second and foreign language classrooms*. London: Continuum.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20(01), 83-108. doi:10.1017/S0272263198001041
- Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46(4), 610-41. doi:10.1002/tesq.34



- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578. doi:10.1093/applin/amp044
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly*, 35(2), 233–49. doi:10.2307/3587647
- Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, 21(2), 45–105.
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *IRAL - International Review of Applied Linguistics in Language Teaching*, 45(3), 193–213. doi:10.1515/iral.2007.009
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor: University of Michigan Press.
- Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.
- Segalowitz, N. (2013). Fluency. In P. Robinson (Ed.), *The Routledge encyclopaedia of second language acquisition* (pp. 240–244). London: Routledge.
- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(2), 173–199. doi:10.1017/S0272263104262027
- Seifoori, Z., & Vahidi, Z. (2012). The impact of fluency strategy training on Iranian EFL learners' speech under online planning conditions. *Language Awareness*, 21(1-2), 101–112. doi:10.1080/09658416.2011.639894
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1–14. doi:10.1017/S026144480200188X
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532. doi:10.1093/applin/amp047
- Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam: John Benjamins.

- Tabachnick, B., & Fidell, L. (1996). *Using multivariate statistics*. London: HarperCollins.
- Tavakoli, P. (forthcoming). Development of fluency in monologic and dialogic task performance. *International Review of Applied Linguistics: Special Issue on Fluency*.
- Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, 65(1): 71-9. doi: 10.1093/elt/ccq020
- Tavakoli, P. & P. Foster (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 58(2): 439-473. doi: 10.1111/j.1467-9922.2008.00446
- Tavakoli, P. & P. Skehan (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language*. (pp. 239-277). Amsterdam, John Benjamins.
- Witton-Davies, G. (2014). *The study of fluency and its development in monologue and dialogue*. Unpublished Doctoral Thesis, University of Lancaster.
- Wright, C. (2013). An investigation of working memory effects on oral grammatical accuracy and fluency in producing questions in English. *TESOL Quarterly*, 47(2), 352–374. doi:10.1002/tesq.68

Table 1: Mean for all fluency measures at Time 1 and Time 2 (standard deviations in parenthesis)

Fluency Measures	Control Group (n=18)		Experimental Group (n=19)	
	Time 1	Time 2	Time 1	Time 2
Mean length of run	6.50 (2.39)	7.51 (3.87)	6.86 (2.78)	8.53 (3.40)
Mean length of silent pauses	.64 (.12)	.69 (.13)	.71 (.18)	.68 (.25)
Phonation time ratio	75.01 (6.68)	74.77 (7.29)	74.14 (7.25)	78.29 (8.67)
Articulation rate	196.34 (27.15)	201.81 (28.95)	188.64 (22.91)	217.73 (18.10)
Speech rate	146.32 (22.01)	152.93 (23.66)	140.51 (24.70)	168.59 (24.99)
Mean number of pauses clause-internal	14.22 (6.32)	13.94 (5.75)	12.84 (4.34)	12.95 (5.57)
Mean number of pauses clause-external	9.39 (3.63)	8.28 (3.10)	8.26 (2.64)	8.21 (2.39)
Mean number of repair measures	7.17 (2.87)	7.44 (2.99)	7.37 (4.11)	5.58 (4.08)
Mean number of filled	10.94	14.44	14.47	14.84

pauses	(4.22)	(6.92)	(5.05)	(6.18)
--------	--------	--------	--------	--------

Table 2: Paired sample t-tests comparing fluency gain scores from Time 1 to Time 2

Fluency Measures	Control Group (n=18)		Experimental Group (n=19)		
	<i>t</i>	<i>P</i>	<i>T</i>	<i>P</i>	<i>Partial</i> <i>η</i> <sup>2</sup>
Mean Length of Run	1.28	.22	2.63	.017*	.18
Mean length of silent pauses	1.52	.15	.65	.53	.00
Phonation time ratio	.20	.84	2.64	.016*	.11
Articulation rate	.64	.53	5.75	.001*	.26
Speech rate	1.13	.27	5.34	.001*	.40
Mean number of pauses clause-internal	.27	.79	.09	.93	.00
Mean number of pauses clause-external	1.72	.10	.07	.94	.03
Mean number of repair measures	.32	.75	1.84	.08	.03
Mean number of filled pauses	2.30	.03	.33	.75	.10

\* Significant at or lower than .025

Table 3: Pearson *r* correlations between speech rate and other measures of fluency

Fluency Measures	Time 1		Time 2	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Mean length of run	.695	.001**	.730	.001**
Mean length of pause	-.493	.002**	-.495	.004**
Phonation time ratio	.539	.001**	.637	.001**
Articulation rate	.727	.001**	.623	.001**
Mean number of pauses clause-internal	-.470	.003**	-.405	.013*
Mean number of pauses clause-external	-.050	.766	-.233	.166
Mean number of repair measures	.227	.176	.242	.149
Mean number of filled pauses	-.013	.940	.217	.197

Significant figures are \*\**p* < .01 and \**p* < .05

Table 4: Means for measures of accuracy and complexity at Time 1 and Time 2 (SD in parenthesis)

Measures	Control Group (n=18)		Experimental Group (n=19)	
	Time 1	Time 2	Time 1	Time 2
Percentage of error-free Clauses	33.65 (15.37)	42.97 (13.45)	36.10 (16.14)	42.12 (18.93)

Percentage of correct use of verb use	79.80 (13.34)	80.60 (15.05)	80.48 (15.70)	82.34 (13.52)
Ratio of subordinate clauses to AS units	1.70 (.30)	1.75 (.35)	1.60 (.31)	1.74 (.35)
Mean length of AS unit	11.71 (1.71)	11.92 (2.15)	12.00 (2.60)	12.60 (2.17)

## Appendix 1

### Monologic task: Time 1

Tell us about the last time you went/did traditional shopping.

You might want to talk about the following:

- What did you do?
- Who did you go with?
- How much time and money did you spend?
- Did you enjoy this experience?

You have 1 minute to prepare and to think about the topic. You can take notes.

### Monologic task: Time 2

Tell us about the first time you arrived at this town and the University.

You might want to talk about the following

- What did you do?
- Where did you stay?
- What was the weather like?
- What were your first impressions of the town/ the university?
- Did you have any difficulties?

You have 1 minute to prepare and to think about the topic. You can take notes.